

A Subset Selection Algorithm for Wireless Sensor Networks

Seyed Hamed Mousavi*, Javad Haghighat*, Walaa Hamouda[†], *Senior Member, IEEE*, and Reza Dastbasteh[‡]

Abstract

One of the main challenges facing wireless sensor networks (WSNs) is the limited power resources available at small sensor nodes. It is therefore desired to reduce the power consumption of sensors while keeping the distortion between the source information and its estimate at the fusion centre (FC) below a specific threshold. In this paper, given the channel state information at the FC, we propose a subset selection algorithm of sensor nodes to reduce the average transmission power of the WSN. We assume the channels between the source and the sensors to be correlated fading channels, modeled by the Gilbert-Elliott model. We show that when these channels are known at the FC, a subset of sensors can be selected by the FC such that the received observations from this subset is sufficient to estimate the source information at the FC while maintaining the distortion between source information and its estimate below a specific threshold. Through analyses, we find the probability distribution of the size of this subset and provide results to evaluate the power efficiency of our proposed algorithm.

Index Terms

Wireless Sensor Networks, Correlated Fading Channels, Gilbert-Elliott model

I. INTRODUCTION

Wireless sensor networks (WSNs) [1], [2], [3], [4] are receiving increasing attention due their numerous current and foreseen applications in several fields including field trials and performance monitoring of solar panels [5], target detection through digital cameras [6], and

*Department of Electrical Engineering, Shiraz University of Technology, Iran (Email: s.moosavi,haghighat@sutech.ac.ir)

[†]Electrical and Computer Engineering, Concordia University, Montreal, QC, Canada, H3G 1M8 (Email: hamouda@ece.concordia.ca)

[‡]Department of Mathematics, Shiraz University, Iran (Email: r.dastbasteh@shirazu.ac.ir)

petrochemical industry fields [7]. One of the main challenges of WSNs is to overcome their energy constraint problem. That is, sensors are powered by batteries with limited energy budgets. Due to deployment of sensor nodes in inaccessible or hostile environments, recharging these batteries is often not an option. Also, the network is expected to have a lifetime in the order of several months, or even years [8]. Therefore, design of power-efficient WSNs is of extreme importance. Numerous works are dedicated to the topic of energy conservation in WSNs. A comprehensive review of these works is given in [8].

A typical sensor node in a WSN consists of three main subsystems namely; sensing, processing, and wireless communication subsystems. These subsystems are responsible for data acquisition, local data processing, and data transmission, respectively [8]. In addition, a power source is included in the sensor node with a limited power budget. In [9], it is shown using experimental measurements that in most cases data transmission is the most energy-consuming unit of the sensor node. In a similar conclusion, it is estimated in [10] that transmitting one bit by the data communication unit requires an energy equivalent to performing about a thousand operations in the data processing unit. It is worth mentioning that in some applications, the sensing subsystem might consume more power than the data communication subsystem (see [8] for details) but in typical applications of WSNs, the highest portion of power is consumed by the data communication subsystem. Therefore, it is highly desired to develop protocols to reduce the transmission power of sensor nodes and hence extend their lifetime.

In this paper, we consider a WSN where the source-sensor channels are correlated fading channels modeled as Gilbert-Elliott channels [11], [12]. The Gilbert-Elliott model is a first-order Markov model for a correlated fading channel quantized to binary levels of Good and Bad states by setting a proper Signal-to-Noise Ratio (SNR) threshold. The channel in its Good and Bad states is modeled by binary-symmetric channels (BSCs) with crossover probabilities of p_G and p_B , respectively. The Gilbert-Elliott model is the simplest possible finite-state Markov (FSM) model for correlated fading channels. The problem of modeling a correlated fading channel by a FSM process is considered in numerous works. An excellent review of works on FSM modeling of fading channels is provided in [13] where the relations between real-valued fading channel parameters and the FSM channel parameters are also considered.

Let a binary source block consisting of M bits be transmitted to N sensors via independent Gilbert-Elliott channels. For each source-sensor channel, the channel states during this transmission can be expressed as an M bit binary sequence where we let a bit 1 represent a Good state

and a bit 0 represent a Bad state. We call this M -bit sequence as the channel-state information Sequence (CSI sequence). For slowly varying fading channels the CSI sequence consists of a few runs and is efficiently compressed using a run-length code (See Fig. 3 and Table I).

Our main contribution in this paper is to propose and analyze a two-phase transmission scheme as follows. At the first phase, each sensor compresses its respective source-sensor CSI using a run-length code and transmits it to the FC. Based on the received CSI from all nodes, the FC will know the location of Good bits, i.e. the bits that are received in a Good channel state. The FC then finds the smallest subset of sensors such that for each source bit, at least one of the sensors in the subset has a Good observation of that source bit. In other words, this is the subset with minimum number of sensors, such that for each source bit at least one of the sensors received this bit in Good channel state. Then, the FC sends a feedback signal to request transmission from this subset. Therefore, at the second phase, only a subset of sensors transmit to the FC, resulting in reduction in the average transmission power.

The motivation behind our proposed algorithm is as follows. According to the Gilbert-Elliott model [11], [12], we have $p_G < p_B$, i.e. Good bits are more reliable than Bad bits. Therefore, we are in fact attempting to find the minimum number of sensors such that if these sensors transmit to the FC and the rest of sensors remain silent, the FC still receives one (or more than one) reliable copy of each source bit and consequently is able to reliably reconstruct the source information. To examine this idea more precisely, let the WSN have an end-to-end distortion requirement of $D \leq \hat{D}$, where D is the expected value of the normalized Hamming distortion (the Bit Error Rate) between the source and its estimate at the FC; and \hat{D} is a fixed distortion threshold. If a (minimum-sized) subset of sensors exists such that each source bit is received through a Good channel by at least one of the sensors in the subset, then the FC will be able to reconstruct the source with a distortion less than or equal p_G . Let ν be the probability of existence of such subset. Then we could bound the end-to-end distortion of the WSN as $D \leq D_u$ where $D_u = \nu \times p_G + (1 - \nu) \times \frac{1}{2}$ where we used the fact that in worst case, the distortion is bounded by $\frac{1}{2}$. In Section V-B, we show that for a WSN with sufficiently large number of sensors, the value of ν is arbitrarily close to 1 and therefore, $\lim_{N \rightarrow \infty} D_u = p_G$. The value of p_G could be expressed as $p_G = \int_{\lambda_t}^{\infty} P_b(\lambda) f(\lambda | \lambda > \lambda_t) d\lambda$ where λ_t is the SNR threshold applied for quantizing the fading channel, $P_b(\lambda)$ is the bit error probability for SNR of λ , and $f(\lambda | \lambda > \lambda_t)$ is the conditional probability distribution function of the SNR. Assuming a binary-phase shift-keying (BPSK) modulation and an additive white Gaussian noise with two-sided power spectral

density of $\frac{N_0}{2}$ at the receiver, we have $P_b(\lambda) = Q(\sqrt{2\lambda})$, where $Q(\cdot)$ represents the Q-function. From the above results, we could bound p_G as $p_G \leq Q(\sqrt{2\lambda_t})$ where for obtaining this upper bound we used the fact that $P_b(\lambda)$ is a decreasing function of λ and has its maximum value at λ_t . In conclusion, for WSNs with sufficiently large number of sensors, the distortion upper bound D_u is always less than or equal to $Q(\sqrt{2\lambda_t})$. Therefore, if λ_t is such that $Q(\sqrt{2\lambda_t}) \leq \hat{D}$, we could conclude that our subset selection algorithm satisfies the distortion requirement of $D \leq \hat{D}$, while reducing the average transmission power of the sensor nodes. In this paper we assume that the condition $p_G \leq \hat{D}$ holds, and proceed with presenting our subset selection algorithm.

The rest of this paper is organized as follows. In Section II we present our system model used in the paper. In Section III, we present our proposed two-phase algorithm with some examples. In Section IV, we analytically derive the probability distribution of the size of the minimum-size subset, as a function of network size, channel parameters, and the source sequence length (the size of this subset is a random variable that depends on the CSI realizations). We also consider the computational complexity of our analytical solution and provide suggestions to reduce this complexity in Section V. In Section VI, we provide numerical results to evaluate the efficiency of our scheme in terms of power conservation. Finally, Section VII concludes the paper.

II. SYSTEM MODEL

We consider a data gathering WSN illustrated in Fig. 1, where an M -bit binary source is sensed by N sensors via Gilbert-Elliott channels and then transmitted to the FC via noiseless channels. To justify the assumption of noiseless sensor-FC channels, we note that according to IEEE 802.15.4 standard, it is recommended that the network combines cyclic redundancy check (CRC) codes with automatic-repeat request (ARQ) and continues re-transmission for a pre-determined number of times [14]. Therefore, assuming genie CRCs, a sensor's data is either eventually delivered to the FC error-free, or not delivered to the FC at all. We assume that the N sensors of Fig. 1 are the sensors that succeeded to deliver their data to the FC before the maximum allowed number of re-transmissions is reached. Also, note that several researchers suggested including a forward-error correction (FEC) scheme at sensor nodes to reduce the error probability of the sensor-FC link (e.g., [15], [16], [17], [18] and references therein). This will reduce the expected number of re-transmission requests.

The state diagram of the Gilbert-Elliott channel is shown in Fig. 2. The channel is modeled by a Good and a Bad states and at each state the channel acts as a BSC with transition probabilities

of p_G and $p_B > p_G$ ($p_G, p_B < 0.5$), respectively. The transition probabilities from the Good state to the Bad state and from the Bad state to the Good state are represented by parameters ϵ and μ respectively. As mentioned in Section I, the Gilbert-Elliott channel can be considered as a quantized version of a correlated fading channel. Figure 3 shows an example of channel realizations for a network with $N = 6$ sensors and $M = 256$ source bits. The dark areas show the Good state and the white areas show the Bad state. To obtain these realizations, we generated 6 realizations of correlated Rayleigh fading channels using Jakes model [19]. Then, we applied a quantization threshold of $\alpha_t = 1$ on the fading amplitude, α . The fading channels have a normalized fading rate of $f_d T_s = 2 \times 10^{-3}$ where f_d is the Doppler frequency and T_s is the symbol period. Through Monte-Carlo simulations, we estimated the resulting Gilbert-Elliott channel state transition probabilities as $\epsilon = 0.0075$ and $\mu = 0.0041$, respectively. It is observed from Fig. 3 that the CSI consists of a few runs and therefore, could be efficiently compressed by a run-length code. In Table I we show the expected value of the compression rate for the run-length coding scheme, for slowly varying fading channels with different normalized fading rates.

III. PROPOSED TWO-PHASE TRANSMISSION ALGORITHM

Assume that we wish to re-construct the binary source at the FC with a normalized Hamming distortion less than or equal to a threshold, \hat{D} . Also assume that $p_G \leq \hat{D}$ and define a *coverage* event as follows:

Definition: A source bit is covered by a subset of sensors if it is sensed via a Good channel by at least one of the sensors in the subset. An M -bit source sequence is covered by a subset of sensors if all of its bits are covered by the subset.

For example, in Fig. 3 the source sequence is covered by the subset consisting of the first, second, and fourth sensors. Given the above definition, our proposed transmission scheme is a two-phase scheme as follows. (i) At the first phase, the sensors transmit their compressed CSI to the FC and then wait for a feedback signal from the FC to proceed. (ii) The FC de-compresses the received CSI and selects the smallest subset of sensors that cover the source sequence. The implementation of the selection algorithm at the FC is shown in Fig. 4. As shown in Fig. 4, if no subset is covering the source sequence, the FC requests transmission from all N sensors, in order to collect all available information for reconstructing the source information. After selecting this minimum size subset, by transmitting a limited feedback (e.g. an N -bit string where the selected

sensors are marked by 1 and the non-selected sensors are marked by 0) the FC informs the sensors of which subset is selected, and only that subset of sensors transmit their observations to the FC. It is clear that receiving observations from this subset is sufficient to recover the source information with a distortion less than or equal to p_G . If p_G is less than or equal to the tolerable distortion threshold of the network, which we represent by \hat{D} , then the received transmissions from the selected subset is sufficient to satisfy the distortion requirement of the system. Also, by applying this subset selection method, only a portion of sensors transmit at each time and therefore, the average transmission power of sensors reduces.

Let us refer to the size of the selected subset by K . Obviously, K is a random variable that depends on the CSI realizations and takes values from 1 to N . The expected value of K is an important indicator in our proposed scheme. The ratio of this expected value to the total number of sensors, N , represents the average ratio of sensors transmitting to the FC. If this ratio becomes smaller, the average transmission power is reduced.

To quantify the power efficiency of our proposed two-phase scheme, we consider the total number of transmitted bits by sensors as an indicator of the consumed power, and compare this parameter with a *conventional* one-phase scheme where all sensors transmit all their observed bits to the FC and no CSI is transmitted. Let us denote the total number of transmitted bits of the conventional scheme and our scheme by B_1 and B_2 , respectively. Obviously we have $B_1 = M \times N$. Also, it is easy to observe that B_2 is a random variable and if the expected value of the compression rate of the run-length coding scheme is represented by $\bar{\rho}$ then, we have $E[B_2] = M \times (\bar{\rho} + E[K])$. Now, if we define an efficiency factor η as the ratio of B_1 and $E[B_2]$, we have:

$$\eta = \frac{N}{\bar{\rho} + E[K]}. \quad (1)$$

If η is greater than one, then our proposed scheme consumes less power compared to the conventional scheme. In Section VI, we evaluate η for Gilbert-Elliott channels with different parameters, as well as for different number of sensors and source sequence lengths, M . Our results show that in many cases, η is considerably larger than one.

IV. PROBABILITY DISTRIBUTION OF THE SELECTED SUBSET SIZE

As mentioned in Section III, the size of the selected subset is a random variable that depends on CSI realizations. Refer to this subset size by K and let $f_K(\cdot)$ and $F_K(\cdot)$ be the probability

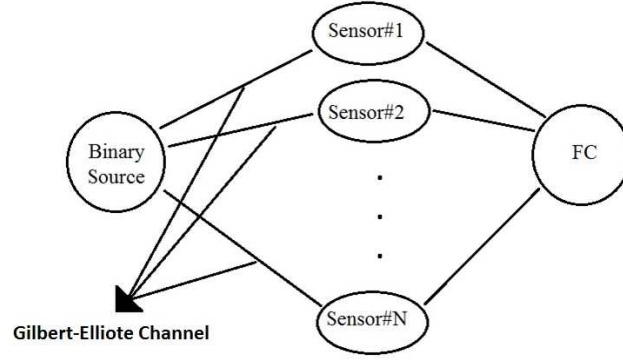


Fig. 1. System model of the data gathering Wireless Sensor Network.

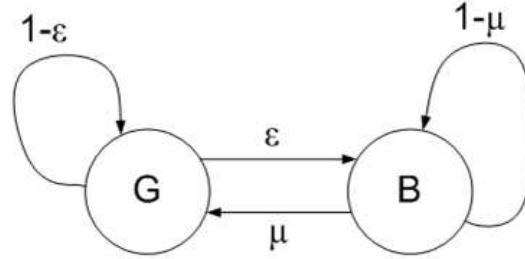


Fig. 2. Gilbert-Elliott channel model

mass function and the cumulative mass function of K , respectively. Obviously, $F_K(k)$ is the probability that there exists a subset of k sensors to cover the source sequence (if a subset of smaller size covers the sequence, we could add arbitrarily selected sensors to this subset to make its size equal to k). Also, the probability mass value, $f_K(k) = F_K(k) - F_K(k-1)$ is the probability that k is the smallest size of a subset that covers the source sequence. In the sequel, we derive an analytical expression for $F_K(k)$.

We assume N independent Gilbert-Elliott channels between source and sensors. Let (μ_n, ϵ_n) represent the state transition probabilities for the channel from the source to the sensor number n . Assume the transmission of bit number m for a fixed m . Let us denote the source-sensor channel states at time interval m by $\mathbf{C}_m = (C_m(1), \dots, C_m(N))$ where $C_m(n) = 1$ if the channel from the source to the n th sensor is in Good state, and $C_m(n) = 0$ if the channel from the source to

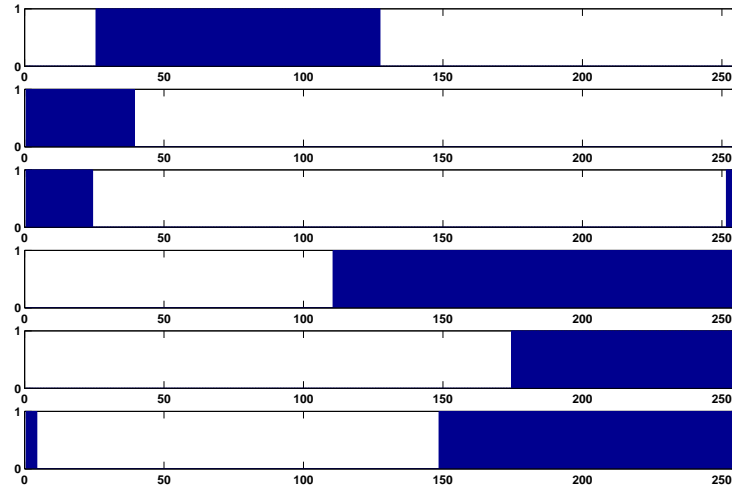


Fig. 3. Realization of Gilbert-Elliott Channel State Information for 6 sensors. The Gilbert-Elliott channels are results of quantizing correlated Rayleigh Fading channels with normalized fading rates of $f_d T_s = 0.002$. A fading amplitude of $\alpha_t = 1$ is used as the threshold for quantization. Dark areas show Good channel states (i.e. amplitudes above the threshold) and white areas show Bad channel states.

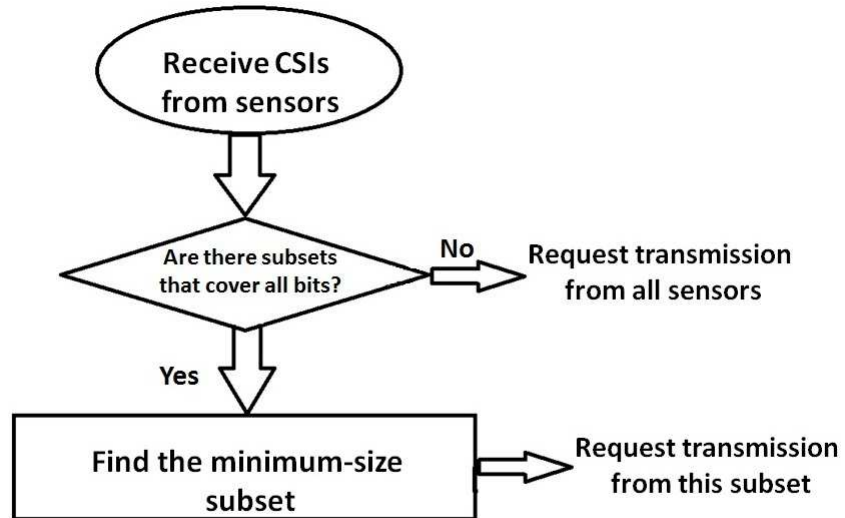


Fig. 4. The sensor selection algorithm at the Fusion Centre.

the n th sensor is in Bad state. Let S be an ordered subset of $(1, 2, \dots, N)$ with cardinality $|S|$ such that $S = (S(1), S(2), \dots, S(|S|))$ and $S(1) < S(2) < \dots < S(|S|)$. Define:

$$\gamma_m(S) = \begin{cases} 1; & \prod_{m'=1}^m \left(\sum_{n=1}^{|S|} C_{m'}(S(n)) \right) > 0 \\ 0; & \text{otherwise.} \end{cases} \quad (2)$$

In (2), $\gamma_m(S) = 1$ if at every bit interval $m' = 1 : m$, the channel state from the source to at least one of the sensors in set S is in Good state, i.e. all m bits are covered by the set S . Note that if at some bit interval m' all these channel states are Bad, then for that bit interval $\sum_{n=1}^{|S|} C_{m'}(S(n)) = 0$ which results in $\gamma_m(S) = 0$.

Using the above definition, $F_K(k)$ is equal to the probability that there exists at least one set S with $|S| \leq k$ such that $\gamma_M(S) = 1$ (M is the total number of transmitted source bits). For calculating this probability, it is sufficient to calculate the probability that there exists a set with $|S| = k$ and $\gamma_M(S) = 1$ (as mentioned above, if a set with cardinality less than k covers all bits up to bit M , we could add a proper number of arbitrarily chosen sensors to make the cardinality of this set k and the extended set still covers all bits up to bit M).

Define $N_k = \binom{N}{k}$. There exist N_k sets S with $|S| = k$ which we refer to as s_1, s_2, \dots, s_{N_k} . Now we can write:

$$F_K(k) = P((\gamma_m(s_1) = 1) \parallel (\gamma_m(s_2) = 1) \parallel \dots \parallel (\gamma_m(s_{N_k}) = 1)) \quad (3)$$

Applying the principle of inclusion and exclusion we have:

$$F_K(k) = \sum_i P(\gamma_m(s_i) = 1) - \sum_{\substack{i, j \\ i \neq j}} P((\gamma_m(s_i) = 1), (\gamma_m(s_j) = 1)) + \dots \quad (4)$$

To simplify the notation, let us define $W = (W(1), W(2), \dots, W(|W|))$ as an ordered subset of $(1, 2, \dots, N_k)$, where $|W| \leq N_k$ is the cardinality of W . Now consider the sets $s_{W(1)}, s_{W(2)}, \dots, s_{W(|W|)}$. It is clear that

$$P((\gamma_m(s_{W(1)}) = 1), (\gamma_m(s_{W(2)}) = 1), \dots, (\gamma_m(s_{W(|W|)}) = 1)) = P\left(\prod_{l=1}^{|W|} \gamma_m(s_{W(l)}) = 1\right).$$

Let us also define:

$$\Gamma_m(W) = \prod_{l=1}^{|W|} \gamma_m(s_{W(l)}). \quad (5)$$

Now, by noting that there are 2^{N_k} possible choices for W , which we represent by $w_1, w_2, \dots, w_{2^{N_k}}$, one can rewrite the inclusion-exclusion expression of (4) as follows:

$$F_K(k) = \sum_{j=1}^{2^{N_k}} (-1)^{|w_j|+1} Pr(\Gamma_m(w_j) = 1). \quad (6)$$

Now let us look at vector \mathbf{C}_m defined above. There are 2^N possible realizations for \mathbf{C}_m which are in fact the 2^N distinct binary n -tuples. We refer to these binary n -tuples by $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{2^N}$. Now the joint probability of events $\Gamma_m(w_j) = 1$ and $\mathbf{C}_m = \mathbf{u}_i$ can be calculated as:

$$P(\Gamma_m(w_j) = 1, \mathbf{C}_m = \mathbf{u}_i) = \sum_{l=1}^{2^N} P(\Gamma_m(w_j) = 1, \mathbf{C}_m = \mathbf{u}_i, \mathbf{C}_{m-1} = \mathbf{u}_l) \quad (7)$$

where we can write:

$$P(\Gamma_m(w_j) = 1, \mathbf{C}_m = \mathbf{u}_i, \mathbf{C}_{m-1} = \mathbf{u}_l) = P(\Gamma_m(w_j) = 1, \Gamma_{m-1}(w_j) = 1, \mathbf{C}_m = \mathbf{u}_i, \mathbf{C}_{m-1} = \mathbf{u}_l). \quad (8)$$

Note that in (8)

$$\begin{aligned} & P(\Gamma_m(w_j) = 1, \Gamma_{m-1}(w_j) = 1, \mathbf{C}_m = \mathbf{u}_i, \mathbf{C}_{m-1} = \mathbf{u}_l) \\ &= P(\Gamma_m(w_j) = 1, \mathbf{C}_m = \mathbf{u}_i, \mathbf{C}_{m-1} = \mathbf{u}_l) \times P(\Gamma_{m-1}(w_j) = 1, \Gamma_m(w_j) = 1, \mathbf{C}_m = \mathbf{u}_i, \mathbf{C}_{m-1} = \mathbf{u}_l) \end{aligned}$$

and $P(\Gamma_{m-1}(w_j) = 1 | \Gamma_m(w_j) = 1) = 1$.

Using (8), one can obtain

$$\begin{aligned} & P(\Gamma_m(w_j) = 1, \mathbf{C}_m = \mathbf{u}_i, \mathbf{C}_{m-1} = \mathbf{u}_l) = \\ & P(\Gamma_{m-1}(w_j) = 1, \mathbf{C}_{m-1} = \mathbf{u}_l) \times P(\Gamma_m(w_j) = 1, \mathbf{C}_m = \mathbf{u}_i | \Gamma_{m-1}(w_j) = 1, \mathbf{C}_{m-1} = \mathbf{u}_l). \end{aligned} \quad (9)$$

Let us rewrite the second term in the righthand side of (9) as follows:

$$\begin{aligned} & P(\Gamma_m(w_j) = 1, \mathbf{C}_m = \mathbf{u}_i | \Gamma_{m-1}(w_j) = 1, \mathbf{C}_{m-1} = \mathbf{u}_l) = \\ & P(\mathbf{C}_m = \mathbf{u}_i | \Gamma_{m-1}(w_j) = 1, \mathbf{C}_{m-1} = \mathbf{u}_l) \times P(\Gamma_m(w_j) = 1 | \mathbf{C}_m = \mathbf{u}_i, \Gamma_{m-1}(w_j) = 1, \mathbf{C}_{m-1} = \mathbf{u}_l). \end{aligned} \quad (10)$$

Note in the righthand side of (10) that given \mathbf{C}_{m-1} , \mathbf{C}_m is independent of $\Gamma_{m-1}(w_j)$. Also given $\Gamma_{m-1}(w_j)$ and \mathbf{C}_m , $\Gamma_m(w_j)$ is independent of \mathbf{C}_{m-1} . This second claim is made by noting that if $\Gamma_{m-1}(w_j) = 1$, then $\Gamma_m(w_j) = 1$ if and only if given the channel realization \mathbf{C}_m , w_j is such that

for every subset $s_{w_j(i)}$, $i = 1 : |w_j|$, at least one of the sensors in the subset has a Good source-sensor channel. Therefore, it is clear that $P(\Gamma_m(w_j) = 1 | \mathbf{C}_m = \mathbf{u}_i, \Gamma_{m-1}(w_j) = 1, \mathbf{C}_{m-1} = \mathbf{u}_l)$ is a function of channel realization \mathbf{u}_i and the set w_j . If we refer to this function by $d_j(i)$, one can write:

$$d_j(i) = \begin{cases} 1; & \prod_{l=1}^{|w_j|} \left(\sum_{h=1}^k \mathbf{u}_i(s_{w_j(l)}(h)) \right) > 0 \\ 0; & \text{otherwise} \end{cases} \quad (11)$$

To clarify this definition, note that if $\sum_{h=1}^k \mathbf{u}_i(s_{w_j(l)}(h))$ is a positive number, then given $\mathbf{C}_m = \mathbf{u}_i$, the subset $s_{w_j(l)}$ covers the m th bit. In fact (11) states that $d_j(i)$ is 1 if for every set $s_{w_j(l)}$, at least one of the sensors in this set receives the m th bit through a Good channel.

Let us define a matrix $Q = [q(i, l)]$ where $q(i, l) = P(\mathbf{C}_m = \mathbf{u}_i | \mathbf{C}_{m-1} = \mathbf{u}_l)$. From the channel model, we can observe that:

$$q(i, l) = \prod_{n=1}^N P(C_m(n) = u_i(n) | C_{m-1}(n) = u_l(n)) \quad (12)$$

and $P(C_m(n) = u_i(n) | C_{m-1}(n) = u_l(n))$ is readily expressed based on the n th source-sensor channel state transition probabilities (μ_n, ϵ_n) .

Now if we define a matrix

$$A_j = [a_j(i, l)],$$

where $a_j(i, l) = d_j(i) q(i, l)$, using (10) and above discussion, one can note that

$$P(\Gamma_m(w_j) = 1, \mathbf{C}_m = \mathbf{u}_i | \Gamma_{m-1}(w_j) = 1, \mathbf{C}_{m-1} = \mathbf{u}_l) = a_j(i, l) \quad (13)$$

and hence (9) can be represented as:

$$P(\Gamma_m(w_j) = 1, \mathbf{C}_m = \mathbf{u}_i, \mathbf{C}_{m-1} = \mathbf{u}_l) = P(\Gamma_{m-1}(w_j) = 1, \mathbf{C}_{m-1} = \mathbf{u}_l) \times a_j(i, l). \quad (14)$$

To simplify (14), let us define a vector

$$\mathbf{X}_m = (X_m(1), X_m(2), \dots, X_m(2^N))$$

where

$$X_m(i) = Pr(\Gamma_m(w_j) = 1, \mathbf{C}_m = \mathbf{u}_i).$$

Now from (7) and (14) we have:

$$X_m(i) = \sum_{l=1}^{2^N} a_j(i, l) X_{m-1}(l) \quad (15)$$

which leads to the following recursive matrix equation:

$$\mathbf{X}_m = A_j \mathbf{X}_{m-1}. \quad (16)$$

Note that to simplify the notation, we dropped dependence of \mathbf{X}_m to j . Also note that A_j is constructed by forcing some rows of matrix Q to zero. Those are the rows i such that $d_j(i) = 0$.

Now from (16), we arrive at the following solution for X_m :

$$\mathbf{X}_m = A_j^{m-1} \mathbf{X}_1 \quad (17)$$

where A_j^{m-1} is the $m - 1$ power of matrix A_j , and the initial vector \mathbf{X}_1 is expressed as:

$$X_1(i) = P(\Gamma_1(w_j) = 1, \mathbf{C}_1 = \mathbf{u}_i) = P(\mathbf{C}_1 = \mathbf{u}_i) P(\Gamma_1(w_j) = 1 | \mathbf{C}_1 = \mathbf{u}_i) \quad (18)$$

It is straightforward to show that

$$P(\Gamma_1(w_j) = 1 | \mathbf{C}_1 = \mathbf{u}_i) = d_j(i). \quad (19)$$

Now noting the independence assumption for source-sensor channels, we have:

$$X_1(i) = d_j(i) \times \prod_{n=1}^N P(C_1(n) = u_i(n)) \quad (20)$$

Following [13] we let the initial channel state $C_1(n)$ have the steady state probability distribution of the corresponding Markov process. For the Markov process of Fig. 2 this steady state distribution is as follows:

$$P(C_1(n) = 1) = \frac{\mu_n}{\epsilon_n + \mu_n} \quad (21)$$

and

$$P(C_1(n) = 0) = \frac{\epsilon_n}{\epsilon_n + \mu_n} \quad (22)$$

After solving (18), we calculate $\Gamma_M(w_j)$ as follows:

$$\Gamma_M(w_j) = \sum_{i=1}^{2^N} X_M(i) \quad (23)$$

and by substituting in (6), we can evaluate $F_K(k)$.

To assess the accuracy of our analyses, in Fig. 5 we compare $F_K(k)$ found using (6) with simulations. For these simulations, 10^5 source sequences of length $M = 128$ bits are transmitted to $N = 5$ sensors via identically distributed Gilbert-Elliott channels with parameters $(\mu_n, \epsilon_n) = (0.0191, 0.0256)$ and 10^5 realizations of K are generated by comparing the 5 corresponding CSIs. The channel parameters (μ_n, ϵ_n) are taken from Table I (see Section VI). It is clear from these results that our analysis is in excellent agreement with the simulated results.

As observed from Fig. 5, $F_K(5) \simeq 0.5$, i.e., in almost 50% of the time, employing all five sensors is not sufficient to cover all source bits. However, as shown in Fig. 4, in these cases, our algorithm forces all sensors to transmit their observations to the FC, i.e. we force $F_K(N) = 1$. In the following section, we will show that by increasing N , the coverage probability increases where the actual values of $F_K(N)$ (before forcing to one) are much closer to one.

Now, the expected value of K can be expressed as:

$$E[K] = \sum_{k=1}^N k \times f_K(k) = N \times F_K(N) - \sum_{k=1}^{N-1} F_K(k)$$

where by noting $F_K(N) = 1$ for our scheme, we reach:

$$E[K] = N - \sum_{k=1}^{N-1} F_K(k). \quad (24)$$

In Section VI, we use the expected value of the subset size, $E[K]$, to evaluate the power reduction achieved by our proposed algorithm.

V. COMPLEXITY AND ASYMPTOTIC PERFORMANCE OF THE PROPOSED ALGORITHM

In what follows, we analyze the computational complexity and asymptotic performance of the proposed two-phase transmission algorithm.

A. Complexity Considerations

Calculating $F_K(k)$ from (6) introduces a computational complexity that is exponentially increasing by N_k where $N_k = \binom{N}{k}$ ¹. Calculation of $F_K(k)$ and consequently, $E[K]$ is

¹Note that this computational complexity only applies to our analysis. Implementing the algorithm at the FC is considerably less complex as in that case the FC has the CSI realizations and only needs to compare them to find the minimum size subset.

time-consuming for large values of N . In fact, the run time for networks with more than $N = 7$ sensors is very large. Therefore, it is desired to introduce bounds on $E[K]$. It is possible to introduce two simple upper bounds on $E[K]$ as follows:

Let $\tilde{F}_K(k)$ be a lower bound for $F_K(k)$. Then, from (24) one can find an upper bound as follows:

$$E[K] \leq N - \sum_{k=1}^{N-1} \tilde{F}_K(k). \quad (25)$$

One possible choice for $\tilde{F}_K(k)$ is by applying Bonferroni's lower bound [20]. Let $L_k \leq N_k/2$ be an integer, then the inclusion-exclusion formula of (6) can be lower-bounded as $F_K(k) \geq \tilde{F}_K(k)$ where

$$\tilde{F}_K(k) = \sum_{\substack{j=1 \\ |w_j| \leq 2L_k}}^{2^{N_k}} (-1)^{|w_j|+1} Pr(\Gamma_m(w_j) = 1). \quad (26)$$

Through simulations, we concluded that for values of L_k which introduce a reasonable computational complexity, the bound of (26) is not tight and in fact leads to a negative value in most cases.

Another simple upper bound can be derived by noting that $F_K(k) \geq F_K(1)$, for $k = 1 : N-1$, which by using (24) leads to:

$$E(K) \leq N - (N-1) F_K(1) \quad (27)$$

where $F_K(1)$ is the probability that one sensor covers the source sequence (i.e., the probability that at least one of the N sensors receives all M source bits via Good source-sensor channels). Fortunately the value of $F_K(1)$ can be simply derived as follows. The probability that the n th sensor covers all source bits equals the probability that the corresponding source-sensor channel is initially at a Good state and stays at this state for the next $M-1$ bit intervals. This probability is equal $\left(\frac{\mu_n}{\mu_n + \epsilon_n}\right) (1 - \epsilon_n)^{M-1}$. Therefore, the probability that none of the N sensors covers the source sequence equal $\prod_{n=1}^N \left(1 - \left(\frac{\mu_n}{\mu_n + \epsilon_n}\right) (1 - \epsilon_n)^{M-1}\right)$ and eventually, the probability that at least one of these N sensors covers the source sequence is given by:

$$F_K(1) = 1 - \prod_{n=1}^N \left(1 - \left(\frac{\mu_n}{\mu_n + \epsilon_n} \right) (1 - \epsilon_n)^{M-1} \right). \quad (28)$$

Note that if all source-sensor channels have identical parameters $(\mu_n, \epsilon_n) = (\mu, \epsilon)$, it is easy to verify that $F_K(1)$ is a monotonically increasing function of N . This is expected, as by increasing the number of sensors, there is a higher probability that at least one of these sensors covers the source sequence.

By replacing $F_K(1)$ from (28) in (27), we find a simple upper bound for $E[K]$ as follows,

$$E(K) \leq N - (N - 1) \left(1 - \prod_{n=1}^N \left(1 - \left(\frac{\mu_n}{\mu_n + \epsilon_n} \right) (1 - \epsilon_n)^{M-1} \right) \right). \quad (29)$$

Figure 6 shows $E[K]$ as a function of N for a network with identical source-sensor channel parameters $(\mu, \epsilon) = (0.0041, 0.0075)$ and source sequence lengths of $M = 200, 256, 300$ bits. The values of (μ, ϵ) are based on Table I. Note the non-monotonic behaviour that is observed in Fig. 6 for the upper bound of $E[K]$. This non-monotonic behaviour is due to two reasons. The first is based on the fact that this upper bound is not tight. However, there is another rational behind the non-monotonic behaviour of this upper bound. That is, in cases where the source is not covered by any subset, we demand transmission from all N sensors (i.e., $K = N$). When N increases, there are subsets with larger sizes to examine for possible coverage. Therefore, $E[K]$ might increase in such cases. Although the upper bound of (29) is not tight, as we will see in Section VI, even by applying this simple bound, we observe considerable power reduction when employing our proposed algorithm for networks with large values of N .

B. Asymptotic Performance

Here, we consider the asymptotic performance of our proposed algorithm for large values of N . For simplicity, let us assume that all N source-sensor channels have identical state transition probabilities (μ, ϵ) . It is clear from (28) that for identical values of $(\mu_n, \epsilon_n) = (\mu, \epsilon)$, $F_K(1)$ is a monotonically increasing function of N and $\lim_{n \rightarrow \infty} F_K(1) = 1$. By noting that $F_K(1) \leq F_K(k) \leq 1$ for $k = 2 : N$, the lower bounds of $\tilde{F}_K(k) = F_K(1)$, $k = 2 : N$ are asymptotically tight. Therefore, the upper bound of (29) is asymptotically tight. If we let $(\mu_n, \epsilon_n) = (\mu, \epsilon)$ and

by taking the derivative of (29) with respect to N , one can find a value N_0 such that for all $N \geq N_0$, the upper bound of $E[K]$ is monotonically decreasing by N

$$N_0 = \left\lceil 1 + \frac{1}{\ln \frac{1}{1-x}} \right\rceil \quad (30)$$

where $x = \left(\frac{\mu}{\mu+\epsilon}\right)(1-\epsilon)^{M-1}$. From the above discussion, we conclude that for sufficiently large values of N , $\frac{E[K]}{N}$ is a monotonically decreasing function of N (decaying by a rate of $\frac{1}{N}$ or faster). Rewrite (1) as

$$\frac{1}{\eta} \leq \frac{\bar{\rho}}{N} + \frac{E[K]}{N} \quad (31)$$

and note that $\bar{\rho} \leq 1$. We observe that $\frac{1}{\eta}$, which is the ratio of power consumption for our proposed algorithm to the conventional transmission scheme, decays by increasing N (at least by a rate of $\frac{1}{N}$). Therefore, our proposed algorithm becomes asymptotically more power efficient by increasing N .

At the end of this section, we note that when we were motivating the idea in Section I, we applied a parameter ν for bounding the distortion, where we defined ν as the probability that there exists a (minimum-size) subset that covers all source bits. We claimed that for sufficiently large N , ν can be arbitrarily close to one. To prove this claim, note that the probability that such subset exists, is greater than or equal the probability that such subset exists and its size is k (for an arbitrarily chosen $k \leq N$). Therefore, $\nu \geq F_K(k) \geq F_K(1)$ and $F_K(1)$ could be arbitrarily close to one, given a sufficiently large N .

VI. NUMERICAL RESULTS

In this section, we provide some numerical results to evaluate the power efficiency of our proposed algorithm. In this work, and without loss of generality, we only consider cases where all source-sensor channels have identical parameters (μ, ϵ) . The Gilbert-Elliott channel parameters are derived by simulating a correlated Rayleigh fading channel using Jakes model and then quantizing the simulated channel by assuming a threshold on the fading amplitude. If we represent the fading amplitude by α and assume that the source is transmitting each bit with energy E_b and the AWGN has a one-sided power spectral density of N_0 , then the instantaneous received SNR equals $\frac{\alpha^2 E_b}{N_0}$ at the sensor. We consider a threshold of $\alpha_{thr} = 1$. That is we assume that SNRs above $\frac{E_b}{N_0}$ leads to a Good delivery of the source bit to the sensor (i.e, the probability of detection error, p_G , is sufficiently low to have $p_G \leq \hat{D}$ as discussed in Section I). The assumption

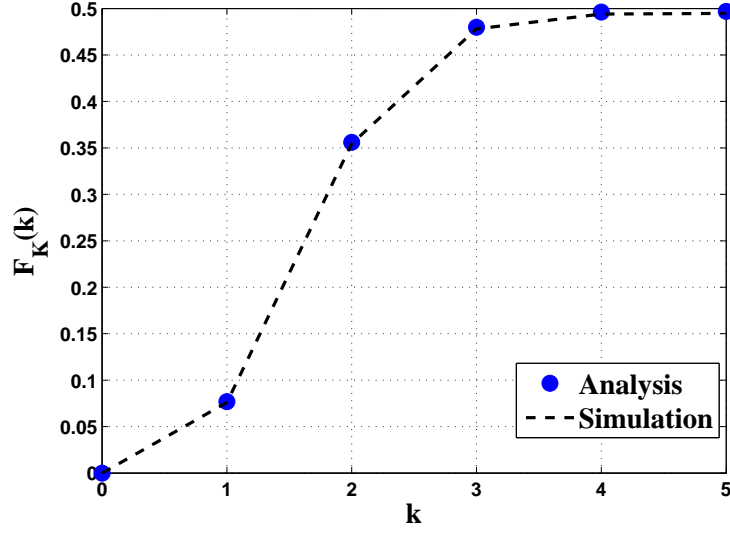


Fig. 5. The cumulative mass function of selected subset size, K , for a network with $N = 5$ sensors and source sequence length of $M = 128$ bits. The source-sensor channels have identical transition probabilities of $(\mu_n, \epsilon_n) = (0.0191, 0.0256)$.

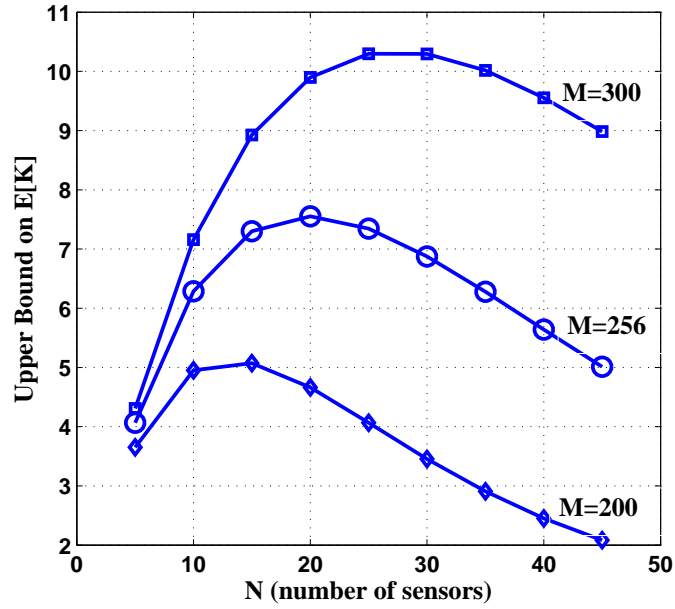


Fig. 6. The upper bound on $E[K]$ for a network with identical source-sensor channel parameters $(\mu, \epsilon) = (0.0041, 0.0075)$. The source sequence lengths are $M = 200$, $M = 256$, and $M = 300$ bits.

of $\alpha_{thr} = 1$ is justified as follows. If we assume that the channel phase shift is perfectly estimated and compensated at the sensor node, then for all $\alpha > \alpha_{thr}$, the channel provides error detection probabilities less than or equal to the error detection probability of an AWGN channel with SNR of $\frac{E_b}{N_0}$. Therefore, by setting this threshold, we eliminate the non-constructive effect of fading and provided source-sensor channels with link qualities equivalent or superior to an AWGN channel. We consider a slow-fading channel, i.e., channels with the normalized fading rates of $f_d T_s \leq 0.01$, where f_d is the maximum Doppler shift and T_s is the symbol duration. The reason we consider slow-fading channels is that as discussed in previous sections, for these channels the run-length coding of CSI sequences provides an efficient compression.

As discussed earlier, the parameters (μ, ϵ) for the Gilbert-Elliott channel are estimated using Monte-Carlo simulation of sufficiently large number of realizations of the fading channel amplitude. For values of $f_d T_s = 0.002, 0.005, 0.008$, the corresponding values of (μ, ϵ) are shown in Table I. Table I also shows the expected value of compression rate for CSIs, for different sequence lengths of $M = 128, M = 256$. As expected, the compression rate decreases when increasing M .

Tables II and III show values of $(E[K], \eta)$ for networks with $N = 4, 5, 6$ sensors. We observe that $E[K]$ is a non-monotonic function of N . The justification of this non-monotonic behaviour was discussed in Section V-B and Fig. 6. Note that the efficiency factor, η , is monotonically increasing function of N , which confirms the increase in efficiency of our proposed algorithm as the number of sensors, N , increases.

From Tables II and III, it is clear that our algorithm is more efficient for channels with slower fading rates. For instance, in Table II, if we let $N = 5$, we observe that the efficiency factor for channels with $f_d T_s = 0.002$ is 1.92 which shows an almost two-fold decrease in power consumption achieved by our algorithm compared to the conventional transmission scheme. However, when we increase $f_d T_s$ to 0.008, η decreases to 1.30. Also, by comparing the results of Table II and Table III, we observe that our proposed algorithm is more efficient for shorter source sequence lengths, M . The reason is that we defined a coverage event as the event that all source bits are covered. Therefore, the coverage probability of a subset decreases by increasing M . This results in an increase of $E[K]$ and consequently a decrease in η . The only case where our scheme shows an inferior performance to the conventional transmission scheme is for $M = 256$, $f_d T_s = 0.008$, and $N = 4$, where $\eta = 0.99$.

To examine the case of large networks with large values of N , we turn to the upper bound

TABLE I

GILBERT-ELLIOTT CHANNEL TRANSITION PROBABILITIES AND ACHIEVED COMPRESSION RATES BY RUN-LENGTH CODING SCHEME FOR DIFFERENT VALUES OF THE NORMALIZED FADING RATE, $f_d T_s$. THE FADING AMPLITUDE THRESHOLD FOR DECIDING BETWEEN GOOD AND BAD STATES IS SET TO 1. THE SOURCE SEQUENCE LENGTHS OF $M = 128$ AND $M = 256$ BITS ARE CONSIDERED.

$f_d T_s$	ϵ	μ	$\bar{\rho}(M = 128)$	$\bar{\rho}(M = 256)$
0.002	0.0075	0.0041	0.1071	0.0813
0.005	0.0165	0.0112	0.1630	0.1454
0.008	0.0256	0.0191	0.2223	0.2134

TABLE II

VALUES OF $(E[K], \eta)$ FOR NETWORKS WITH $N = 4, 5, 6$ SENSORS AND SOURCE SEQUENCE LENGTH OF $M = 128$ BITS.

$f_d T_s$		$N = 4$	$N = 5$	$N = 6$
0.002		(2.44, 1.57)	(2.49, 1.92)	(2.43, 2.37)
0.005		(2.97, 1.28)	(3.12, 1.52)	(3.05, 1.87)
0.008		(3.33, 1.13)	(3.63, 1.30)	(3.55, 1.59)

of (29). Replacing this upper bound in (1) provides a lower bound on η . Figure 7 shows the upper bound of $E[K]$ and the resulting lower bound on η for networks with source sequence length of $M = 256$ bits and $f_d T_s = 0.002$. One can note the considerable gains for these large values of N when using our algorithm. For example, for a network with $N = 50$ sensors, our proposed algorithm provides at least a twelve-fold decrease in the consumed power compared to the conventional transmission scheme with all nodes transmitting ($\eta > 12$). To examine the effect of different block lengths on η , we also consider block lengths $M = 200$ and $M = 300$ in Fig. 7. As observed, the efficiency factor decreases by increasing the block length. This is due the fact that as M increases, the probability that k sensors cover all M bits decreases. As a result $E[K]$ increases and η becomes smaller. Nonetheless, we observe that for $M = 300$ and $N = 50$, our algorithm has an efficiency factor close to 6.

VII. CONCLUSION

We analyzed a WSN where source-sensor channels are modeled as quantized correlated fading channels (Gilbert-Elliott channels). We proposed a two-phase transmission scheme where at the first phase compressed channel state information sequences are transmitted to the FC and a

TABLE III

VALUES OF $(E[K], \eta)$ FOR NETWORKS WITH $N = 4, 5, 6$ SENSORS AND SOURCE SEQUENCE LENGTH OF $M = 256$ BITS.

$f_d T_s$	$N = 4$	$N = 5$	$N = 6$
0.002	(3.06, 1.27)	(3.25, 1.50)	(3.31, 1.77)
0.005	(3.62, 1.06)	(4.05, 1.19)	(4.23, 1.37)
0.008	(3.83, 0.99)	(4.46, 1.07)	(4.78, 1.20)

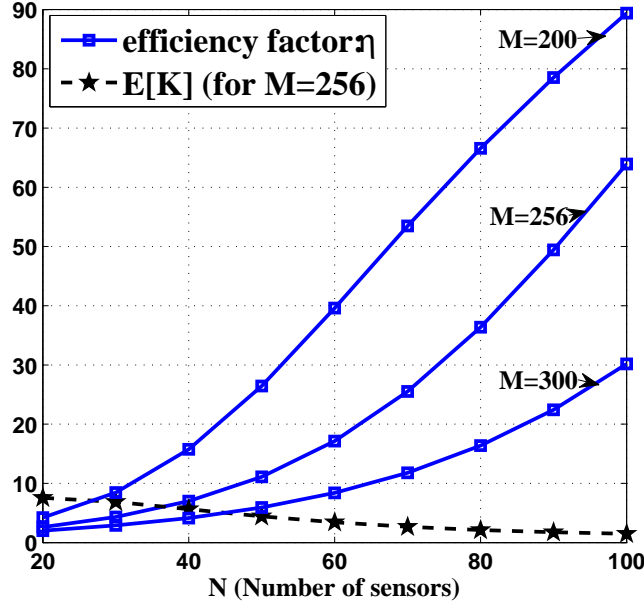


Fig. 7. Values of η for networks with source sequence lengths of $M = 200, 256, 300$ bits and $f_d T_s = 0.002$. The dashed line shows $E[K]$ for $M = 256$ bits.

subset of sensors are selected to transmit their observations to the FC at the second phase. Also, we analytically derived the probability distribution of the size of the selected subset and the expected value of this subset size. We presented simulation results to assess the accuracy of our analyses. We defined an efficiency factor for our proposed algorithm and evaluated this factor for several channel conditions and network setups. In most cases our proposed two-phase algorithm showed a superior power efficiency compared to a conventional one-phase transmission scheme over slow-fading channels.

REFERENCES

- [1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Communications Magazine* (2002), Pages: 102-114.
- [2] Chin-Liang Wang, Syue-Ju Syue, "An Efficient Relay Selection Protocol for Cooperative Wireless Sensor Networks," *Wireless Communications and Networking Conference*, 2009. WCNC 2009. pp.1-5, April 2009.
- [3] Zorzi, M.; Rao, R.R., "Geographic random forwarding (GeRaF) for ad hoc and sensor networks: multihop performance," *Mobile Computing, IEEE Transactions on* , vol.2, no.4, pp.337,348, Oct.-Dec. 2003
- [4] X.J. Zhang and Y. Gong, "Joint power allocation and relay positioning in multi-relay cooperative systems", *IET Commun.* (2009), vol. 3, Issue: 10, Pages:1683-1692.
- [5] Ranhotigamage, C., Mukhopadhyay, S.C., "Field Trials and Performance Monitoring of Distributed Solar Panels Using a Low-Cost Wireless Sensors Network for Domestic Applications," *Sensors Journal, IEEE* , vol.11, no.10, pp.2583,2590, Oct. 2011
- [6] Mahmut Karakaya, Huirong Qi, "Target detection and counting using a progressive certainty map in distributed visual sensor networks" *International Conference on Distributed Smart Cameras, ICDSC 2009*, Pages: 1-8.
- [7] Zhang Ke, Li Yang, Xliao Wang-hui, Suh Heejong, "The Application of a Wireless Sensor Network Design Based on ZigBee in Petrochemical Industry Field" *International Conference on Intelligent Networks and Intelligent Systems, ICINIS 2008*, Pages: 284-287.
- [8] G. Anastasi, M. Conti, M. Di Francesco, A. Passarella, "Energy conservation in Wireless Sensor Networks: A Survey," *Elsevier Ad Hoc Networks*, 7 (2009) 537-568.
- [9] V. Raghunathan, C. Schurghers, S. Park, M. Srivastava, "Energy-aware wireless microsensor networks," *IEEE Signal Processing Magazine* (2002) Pages: 40–50.
- [10] G. Pottie, W. Kaiser, "Wireless integrated network sensors," *Communication of ACM* 43 (5) (2000) 51–58.
- [11] E.N. Gilbert, "Capacity of a burst-noise channel," *Bell Syst. Tech. J.*, vol. 39, no. 9, pp. 1253–1265, Sept. 1960.
- [12] E.O. Elliott, "Estimates of error rates for codes on burst-noise channels," *Bell Syst. Tech. J.*, vol. 42, no. 5, pp. 1977–1997, Sept. 1963.
- [13] P. Sadeghi, R. A. Kennedy, P. B. Rapajic, and R. Shams, "Finite-State Markov Modeling of Fading Channels," *IEEE Signal Processing Magazine*, Sep. 2008, pp. 57-80.
- [14] A. Willing, "Recent and emerging topics in wireless industrial communications: A selection," *IEEE Transactions on Industrial Informatics*, vol. 4, n0. 2, pp. 102-124, May 2008.
- [15] Sheryl L. Howard, Christian Schlegel, and Kris Iniewski, "Error control coding in low-power wireless sensor networks: When is ECC energy efficient," *EURASIP Journal on Wireless Communications and Networking*, pages 1–14, 2006.
- [16] M. R. Islam "Errorcorrection codes in wireless sensor network: An energy aware approach," *Int. J. Comput.Inf. Eng.*, vol. 4, no. 1, pp.59 -64 2010
- [17] Abedi, A., "Power-efficient-coded architecture for distributed wireless sensing," *Wireless Sensor Systems, IET* , vol.1, no.3, pp.129,136, September 2011
- [18] J K. Yu, F. Baracy, M. Gidlundz and J. Akerbergz, "Adaptive forward error correction for best effort wireless sensor networks," *IEEE International Conf. on Communications (ICC)*, Ottawa ,June 2012, pp. 7104-7109.
- [19] W. C. Jakes, *Microwave Mobile Communications*, New York: IEEE Press, 1974.
- [20] C. Bonferroni, "Teoria statistica delle classi e calcolo delle probabilità," *Pubbl. d. R. Ist. Super. di Sci. Econom. e Commerciali di Firenze* (in Italian), vol. 8, Pages: 1–62, 1936.